PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

WORKING PAPER N° 2013 – 04

Where the Streets Have a Name: Income Comparisons in the US

Abel Brodeur
Sarah Flèche

JEL Codes : C25, D00, J31

Keywords: Income comparisons; Rank; Relative utility; Social interactions; Social staut; Well-being

PARIS-JOURDAN SCIENCES ECONOMIQUES
48, BD JOURDAN – E.N.S. – 75014 PARIS
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10
www.pse.ens.fr

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE – ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES
ÉCOLE DES PONTS PARISTECH – ECOLE NORMALE SUPÉRIEURE – INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE

Where the Streets Have a Name: Income Comparisons in the US

Abel Brodeur*and Sarah Fleche†

February 26, 2013

Abstract

This paper analyses how neighbors' income affect agents' well-being using unprecedented data from the BRFSS and the City of Somerville. We conduct a multi-scale approach at the county, ZIP code and street-levels and find that the association between well-being and neighbors' income follows an inverted U-shaped pattern in the size of the area. We find a negative relationship between well-being and neighbors' income in the county of residence, but the opposite at the ZIP code-level. Our results are consistent with the fact that agents enjoy living in a rich ZIP code but also having poor faraway neighbors since they have preferences for high social status. We test explicitly this interpretation by including amenities and the relative rank in the local income distribution in our model. At the street-level, we find a negative association between neighbors' income and self-reported well-being indicating the presence of income comparisons between very close neighbors.

JEL: C25, D00, J31

Keywords: Income comparisons, rank, relative utility, social interactions, social status, well-being

*Paris School of Economics, CEP (LSE) and IZA, 48 Boulevard Jourdan, 75014 Paris, +33(0)1 43 13 63 06. E-mail: abel.brodeur@parisschoolofeconomics.eu (corresponding author).

†Paris School of Economics and CEP (LSE). E-mail: sarah.fleche@ens.fr. We are grateful to Clément Bosquet, Andrew Clark, Jan-Emmanuel De Neve, Xavier Fontaine, Pierre-Yves Geoffard, Erzo Luttmer, Thomas Piketty, Steve Pischke, Nattavudh Powdthavee, Claudia Senik, Malgorzata Switek and Ekaterina Zhuravskaya for valuable comments. Thanks to seminar participants at the 2012 ISQOLS Conference, Jamboree EDP, LSE and PSE. We are also grateful to Judy Bass, Brandi L. Bennett, Michelle L. Cook, Joseph Grandpre, Mark Gildemaster, Kip Neale, Samara Viner-Brown and Jennifer Wrathall, BRFSS coordinators, for valuable comments and providing ZIP codes for their respective state. We thank Daniel Hadley, Peter Heller and the City of Somerville who answered our questions and provided data. Any remaining errors are our own. Financial support from the Fonds Québécois de la Recherche sur la Société et la Culture is gratefully acknowledged by Abel Brodeur.

Average Joes always envy rich people. This reflects the broad idea that we compare to people with whom we interact. This group of agents may include people having the same sex, age or education-level. But, the simplest idea is that we compare to individuals with whom we live, our neighbors. Our question of interest is how do neighbors' income affect well-being? This relationship has important implications. If individuals derive utility from neighbors' earnings, then shocks to one's neighbors' may affect one's consumption. The economic analysis of such relative income effects can be traced back to at least Veblen (1899) and Duesenberry (1949). In terms of policy, tax hikes may change these externalities by lowering or increasing the relative income gap (Boskin and Sheshinski (1978)). Public spending and housing policies also play a critical role in changing local area characteristics and modifying income spillovers.

However, whether neighbors' income has a negative or a positive impact on people's well-being is not straightforward. Several channels may be identified. A first channel is income comparisons. When agents in our place of residence earn more, we feel worse off because it lowers our relative income. This may also capture an element of social status. Agents may enjoy having poorer faraway neighbors since they have preferences for high social status. A second channel is amenities. Neighbors' income is correlated with local area characteristics. Richer areas have on average less criminality, better schools and a good economic environment. Therefore, agents may enjoy living with rich people because they pay more taxes and provide better amenities. Other channels may include a tunnel effect. Neighbors' income contain information about our own future prospects. Overall, the net effect of relative income may be negative or positive depending on the relative size of these channels.

The literature exhibits no consensus about the effect of neighbors' income on well-being. Blanchflower and Oswald (2004), Helliwell and Huang (2011) and Luttmer (2005) report that subjective well-being is positively associated with own income and negatively correlated with average/median income in the region of residence. They argue that the negative effect of neighbors' earnings on well-being is due to relative consumption. Luttmer (2005) used the 1987-88 and 1992-94 waves of the National Survey of Families and Households and matched this data to the Public

Use Microdata Areas ("PUMAs"). These areas have about 150,000 inhabitants on average. Blanchflower and Oswald (2004) replicated a similar analysis using the General Social Survey and state income per capita while Helliwell and Huang (2011) relied on county-level data. They limit their analysis on income comparisons by confirming the findings of Luttmer (2005).

Other studies rely on more disaggregated data and find opposite results. Clark et al. (2009) and Dittmann and Goebel (2010)[1] use respectively Danish and German data. In their papers, they report that agents are more satisfied when their neighbors are richer which contradicts the studies cited above and is consistent with a public good interpretation. While it is possible that Americans and Europeans are differently affected by income comparisons, our intuition is that the disparities in those results are driven by the size of the local area.

Last, Kingdon and Knight (2007) use 366 randomly selected clusters covering 2,900 people on average and broader districts in South Africa. Their findings indicate a positive relationship for clusters and a negative association between neighbors' income and household satisfaction at the district-level. They note that most of the clusters are racially homogeneous and that their results suggest evidence of empathy for close neighbors and comparisons for those further away. In an unpublished paper, Barrington-Leigh and Helliwell (2008) rely on different Canadian surveys and show the relationship between measures of well-being and income spillovers for different geographic scales. Their findings suggest that the overall spillover effect is negative.

Average and median income within a very large area may capture other effects than income comparisons such as unemployment rate, criminality and public spendings. The lack of large data sets and finely disaggregated data make the identification of the different channels difficult. Our paper extends the investigation of this literature on income comparisons and well-being in several ways. We conduct a multi-scale approach at the county and ZIP code-levels, using a very large data set, the Behavioral Risk Factor Surveillance System (BRFSS), and administrative data. We also

[1]Dittmann and Goebel (2010) measure the income of neighbors using an index covering information such as the level of occupation, purchasing power and the rate of self-employment. See also Knies et al. (2008) for a similar analysis using German ZIP codes covering 9,000 inhabitants on average. Last, Clark et al. (2009) rely on geographical grid of 10 000 square meters.

explicitly test the different mechanisms behind the association between neighbors' income and well-being.

Our main findings indicate that faraway and close neighbors' income affect differently agents' well-being. We include both the median income in the county and the ZIP code of residence in our model and find that individuals report higher levels of well-being when their close neighbors (ZIP code) are rich and the opposite for the county of residence. This confirms the idea that there is a stronger identification to close neighbors. We tend to care more about individuals close to us and less about those in a faraway city. Moreover, city neighbors share the same public goods and are potential coworkers. While there still might be a negative effect of comparisons, it seems that it is overwhelmed by a public goods effect when the analysis is at the ZIP code-level. Interestingly, both poorer and richer respondents are affected similarly by the positive effect of income spillovers from close neighbors, but this is not the case at the county-level. The negative association is statistically stronger for residents poorer than the ZIP code median income which could mean that both types of agents benefit from amenities but that social status matters more for poorer residents.[2]

Our intuition of the findings is confirmed by controlling for a large set of amenities in the local area. Adding amenities in our model increases economically and statistically significantly the negative impact of neighbors' income on respondents' satisfaction. This confirms our results that agents enjoy income spillovers because of the public goods. Interestingly, local area characteristics such as criminality and economic activity seem to affect more the variable county median income than direct expenditures by the local government. This might be an indication that agents prefer to live with rich neighbors mainly because it provides a safer environment and better economic conditions. We also present specifications where we add the relative position in the local income distribution. Conditional on own income and relative income rank in the county, the relationship between life satisfaction and median county income is now positive and significant.

[2]We find that individuals are sensitive to poverty since they prefer living in counties where they are among the richest people, but also where poverty is less prevalent.

In the last section, we rely on an unprecedented survey conducted by the City of Somerville, Massachusetts. This survey contains information on self-reported well-being, income and address of residence. We find evidence that residents of Somerville compare to people living in their street. At this very finely disaggregated level, income comparisons seem to dominate the impacts of other channels. Our results show that, conditional on own income, respondents report lower levels of happiness when their next door neighbors are richer. On the other hand, median income of neighbors living in Somerville but not very close to the respondent is positively related to respondents' well-being but statistically insignificant. We believe using such a disaggregated dataset gives a first piece of evidence of the presence of income comparisons between next door neighbors.

Our multi-scale approach at the county, ZIP code and street-levels support the idea that the association between well-being and neighbors' income follows an inverted U-shaped pattern in the size of the area. The findings are consistent with a model where income comparisons, social status and amenities affect agents' well-being. At the county-level, the negative impact of social status on well-being dominates the positive effect of public goods whereas it is the opposite at the ZIP code-level. We also find evidence that residents of Somerville compare to people living in their street. Next door neighbors share similar public goods and are more likely to know the consumption of the inhabitants of their neighborhood.

Next section reviews the literature on relative utility and provides a theoretical framework. Section II details the data at the county and ZIP code-levels. The third section presents the findings using the BRFSS and administrative data while section IV uses a survey collected by the city of Somerville. Section V concludes.

I. Utility and Income Comparisons

The effect of neighbors' income on utility is, *a priori*, ambiguous since many channels are at work. In this section, we present different mechanisms which could explain the relationship between income of neighbors and self-reported well-being. We focus mainly on (i) income comparisons and social status and (ii) local public goods since

we test those explicitly in the next sections.[3]

(i) Income comparisons: well-being depends partly on individual's absolute income and partly on individual's relative income, in the sense that an agent's well-being depends on the gap between own income and some reference benchmark. Getting a new car seems essential when one of our neighbors just bought one (Kuhn et al. (2011)). We may expect that people are worse-off when agents in their reference group do better. This refers to a "jealousy" effect. (see Veblen (1899), Duesenberry (1949) and Stigler (1950)).

Traditionally, this is tested through the coefficient of y^*, the mean or median reference group income. In our contextual framework, y^* is the median income of neighbors. The following relation is assumed:

$$U_i = U(u_1(y_i), u_2(y^*), (X)) \tag{1}$$

where U is the economic concept of utility which depends on y the household income and y^* the place of residence median income. X is a set of individual covariates.

Our econometric model is as follows:

$$SWB_{ijt} = \alpha + \delta ln(y_{it}) + \theta ln(y^*_{jt}) + \gamma X_{it} + \varepsilon_{ijt} \tag{2}$$

where SWB is the outcome variable (for instance: life satisfaction) for respondent i in year t living in j and δ is the coefficient associated with household income.[4] The utility function is believed to be concave in household income which explains our choice to introduce income in logarithmic form. We also rule out the concern that income spillovers proxy for nonlinearities in the impact of own household income by using household income dummies instead of computing the "Log of Household Income" in some specifications. This has no effect on the findings presented in this

[3]Another economic channel is inflation. Imagine geographical locations with higher income have more inflation. This effect is partially captured by the introduction of the median value of housing units in the model.

[4]There is by now a substantial empirical literature regarding such relative utility effect (Card et al. (2012); Clark and Oswald (1996); Easterlin (1995); Frank (1985); Kapteyn et al. (1985); Robson (1992); VanPraag (1971)).

paper.

θ is the coefficient of interest in this framework. We compute cell averages in order to measure y^*. The cells are created using an external data set at the county and ZIP code-levels.[5] We follow the literature by using median income household since it is less sensitive to outliers than the mean (Clark et al. (2008)).

This specification has been used in previous studies using very large area to compute the median income. We believe doing such an exercise requires very small areas since comparison is especially strong with close neighbors. It is also plausible that other mechanisms not related to comparisons affect utility through the median income variable. For instance, median income could be an indicator of social class or future income of young respondents.

We try to disentangle some of these different mechanisms by introducing the relative position of the respondent in the income distribution of the place of residence. Our prediction is that the relative rank in the local income distribution affects positively utility.

More specifically, our econometric model is the following:

$$SWB_{ijt} = \alpha + \delta ln(y_{it}) + \theta ln(y^*_{jt}) + \mu f_{ijt} + \gamma X_{it} + \varepsilon_{ijt} \tag{3}$$

with

$$f_{ijt} = \frac{i-1}{n-1} \tag{4}$$

where n is the number of individuals in the reference group and i the respondent's position.

Appendix D describes how our measure of ranking, f_{ijt}, is constructed. Intuitively, a ranking measures the position of an individual in a specific group (see Brown et al. (2008) and Powdthavee (2009)). In our case, we attribute to respondents a relative rank in the distribution of household income for each county/ZIP code. This normalized rank is equal to zero for the poorest household and one for

[5]Another methodology used in this literature is to compute a predicted income based on individuals characteristics such as place of residence, age, race or gender. See Clark et al. (2008), DiTella and MacCulloch (2006) and Frey and Stutzer (2002) for an aperçu of the relevant papers in this literature.

the richest household in the county/ZIP code.

Including the relative position in the income distribution is interesting for many reasons. First, it allows us to verify if agents care about how many individuals in their region of residence are relatively richer or if they care about how much they are relatively richer. In other words, utility may depend on the importance of the relative position in society or on the income gap between you and your neighbors. Second, median income and rank should both capture relative income. Once we introduce the rank in the local income distribution, it is likely that the median income will show the impact of other channels such as public goods.

(ii) Local Public Goods: we will test explicitly the public good interpretation by including many amenities and county-level characteristics in our model. Wealthy neighborhoods have usually better schools and nicer parks. It may be the case that we want to have rich neighbors, not because they are rich *per se*, but because they pay more taxes and provide better amenities. This may also work through criminality and unemployment. If neighbors' income is positively related to utility, then controlling for public goods should decrease the size of the coefficient of interest (y^*). On the other hand, if there is a negative relationship between utility and median household income, then taking into account public goods should make the coefficient more negative. Recall that in the US, many public goods are provided by the states and counties (education, health care, roads...). The complete list of local area variables used in this study may be found in Appendix B and C.

We test explicitly the public goods, social status and income comparisons channels in the next sections using data from the BRFSS (counties and ZIP codes) and Somerville's survey (streets). We will analyze theories of relative utility using different geographical areas and test whether the income comparison process is scale sensitive. Furthermore, we test the presence of nonlinearities and heterogeneity. It may be the case that poorer and younger individuals are affected by income of their neighbors, while richer and older individuals are not.

(iii) Other channels: many other channels may also affect the association between neighbors' income and well-being. Median income contains potentially information about own social status, but also future prospects. Younger agents may thus compare

more than older agents. This could indicate that there is a "tunnel" effect. Hirschman and Rothschild (1973) put forward the idea that societies experiencing economic development may show, at the beginning of the process, more tolerance towards inequality. The relative increase of others' income is seen as promising evidence about the individual's own chance of success. This first step is then followed, if beliefs are not met, by a decrease in tolerance of inequality (see Grosfeld and Senik (2010) for empirical evidence in Poland). Another important element is social distance. We tend to care more about individuals close to us and less about those in a faraway city. This could mean that income of close neighbors would affect us in a different way than income of faraway neighbors. Residents of your city (ZIP code) have a similar social status or at least one closer than individuals in your county. There is thus a stronger identification to those neighbors which could be interpreted as a signal of social class.

We are aware that the association between well-being and neighbors' income may be spurious due to omitted socioeconomic and local variables. We include individual covariates in our models and will verify whether including housing prices and a broad set of local area characteristics affect our estimates. Next sections will review the complete lists of variables included in our specifications.

II. Behavioral Risk Factor Surveillance System

A. BRFSS: County

This paper is based on data from different surveys. We first rely on the Behavioral Risk Factor Surveillance System (BRFSS) which was established in 1984 by the Centers for Disease Control and Prevention (CDC) but did not include a question on life satisfaction before 2005 (Brodeur (2012); Goudie et al. (2011); Helliwell and Huang (2011); Oswald and Wu (2010)). The time period covered with this data set is thus 2005-2010. The BRFSS is repeated cross section, has a total sample of around 1,750,000 and contains information on county of residence, household income and life satisfaction. It covers more than two thirds of the counties in the US: county

codes are suppressed for counties with fewer than 10,000 residents for confidentiality reasons and statistical reliability.[6]

The following question is asked over the period 2005-2010 in the BRFSS: "In general, how satisfied are you with your life?" where respondents have 4 choices (4=very satisfied, 3=satisfied, 2=dissatisfied and 1=very dissatisfied). Table 1 presents means and standard deviations of the variables coming from the BRFSS and shows the distribution of life satisfaction. 45% of the respondents reported that they were very satisfied with their life. On the other hand, 1% answered that they were very dissatisfied. The question on household income is the following: "Is your annual household income from all sources" where respondents have 8 different choices going from "Less than 10,000" to "75,000 or more". Respectively 4,8% and 32,1% of the individuals report having less than $10,000 and more than $75,000. We divide/multiply bottom/top-coded categories by a factor in order to have an income distribution closer to real figures. We use different factors throughout this paper in order to test the robustness of our results (see Appendix A for more details).

We match the county of residence of respondents in the BRFSS to county-level variables. There are on average 62 counties per state. The states with respectively the smallest and highest number of counties are Delaware (3) and Texas (254). The median land area of counties is on average 622 square miles for the US. County-level data used in this paper come from the U.S. Census Bureau, USA Counties, website.[7] USA Counties collects thousands of data items from a variety of sources such as the Bureau of Economic Analysis, the Department of Education, the Federal Bureau of Investigation and from the 2010 Census of Population and Housing. Appendix B gives a definition of the county-level variables used in the analysis coming from this source: unemployment rate, number of violent crimes known by police, local government direct general expenditures per capita...

Our main variable coming from this data set is the county median household income over the period 2005-2009. Since there is no information on median household income in 2010, the last year available (2009) is used as a replacement. Using the

[6]States have different rules for the data files. It seems though that states report county/ZIP code-level data when the number of respondents is greater than fifty in a given geographic location.

[7]http://censtats.census.gov/usa/usa.shtml.

2010 wave has no effect on the findings reported in this paper. Additionally, USA Counties gives the percentage of people of all ages in poverty, the percentage of related children age 5 to 17 in families in poverty and the percentage of people under age 18 in poverty.

B. BRFSS: ZIP Codes

We also rely on ZIP code of residence. Since the public use version does not identify the ZIP code of residence, we obtained this information from the BRFSS state coordinators. We managed to cumulate the ZIP codes for respondents of 8 states: Arizona, Maine, Ohio, Rhode Island, South Dakota, Texas, Utah and Wyoming.[8] Fortunately, there is at least one state per region (Northeast, Midwest, South and West). The period covered is 2005-2010 for all these states except Texas (2007-2010). There are 33,120 five-digit ZCTAs and 3,033 county-equivalent in the US as of 2010. Summary statistics for the 8 states are presented in Appendix Table 1.

We follow the recommendation of BRFSS coordinators and restrict the sample to ZIP codes where the number of respondents is greater than 50 for statistical reliability. We combine all the years when doing such an exercise which increases the number of ZIP codes that we may use. This gives us a sample size of 221,110 respondents. This technique gives us respectively 167, 212, 345, 59, 156, 269, 128 and 81 ZIP codes for Arizona, Maine, Ohio, Rhode Island, South Dakota, Texas, Utah and Wyoming. We also, for some specifications, restrict the sample to ZIP codes having more than 50 respondents for a given year which gives us a sample size of 119,141. Last, we obtain the ZIP code median household income for 1999 and 2011 from the U.S. Census Bureau. Estimates throughout rely on the 2011 values, but our main results are strikingly similar with data from the 2000 Census.

[8]The remaining states were excluded for three reasons. First, some states did not answer our request. The second reason is simply that many states refused to provide the data in order to protect the confidentiality of respondents. Lastly, we did not have the funding to pay for the fees asked by few states.

III. County and ZIP Code

A. County-level

This section analyzes the relationship between utility of respondents and income of their reference group. As mentioned before, utility is measured by self-reported well-being and the reference group is composed of individuals living in the same geographical area.[9] We first use counties of residence as our unit of social comparisons and then extend our analysis to local data. We exploit the large sample of the BRFSS and match the county of residence with the variables of USA Counties. Our methodology has the advantage of combining a rich survey and reliable administrative data.

Table 2 presents our estimates using county-level data. Our econometrics model is similar to equation (2) presented previously. Life satisfaction is the outcome variable for individual i in county j in year t and state and year fixed effects are included. Our variable of interest is median household income. We rely on Ordered Probit throughout this paper but obtain similar findings with OLS (available upon request). The first column shows the relationship between median household income at the county-level and respondents' well-being without controlling for socioeconomic characteristics. As expected, the log of own household income is positive and statistically significant at the 1 % level. The coefficient of our variable of interest "Log of Median Real Household Income at the County-Level" is negative and statistically significant which means that conditional on own income, respondents report lower levels of satisfaction when their neighbors are richer. This is consistent with previous findings reported by Luttmer (2005). Let's also note that the coefficient of our variable of interest is much smaller than the coefficient of "Log of Real Household Income" indicating that own income is more important than relative income. The estimates imply that if both own household income and neighbors' household

[9]Knight et al. (2009) note that two thirds of respondents in a survey of Chinese households report that their main comparison group consists of individual in their own village. On the other hand, Clark and Senik (2010) find that the large majority of Europeans compare to their work colleagues and also to their friends.

incomes rise by the same percentage, a person would be more satisfied.[10]

Column 2 adds individual covariates to this basic model. We include the following variables: age, age-squared, gender, 8 dummies of working status (employed, unemployed for less than a year, ...), 5 education dummies, 6 dummies of marital status, 4 child dummies and 7 race dummies. While we do not display the coefficients for these variables, they attract signs that are consistent with those of the literature. For instance, there is a positive relationship between life satisfaction and being employed, married or white. Including these controls decreases the size of the coefficients of interest which mean that socioeconomic characteristics are correlated with own income and neighbors' income. Let's note that the regressions include different proxies of income such as marital status, education-level and employment status. It is likely that respondents' socioeconomic variables are correlated to local area characteristics. Better-educated individuals tend to live together for instance.

Column 3 verifies the robustness of these findings by replacing the "Log of Real Household Income" by the "Log of Real Household Income per Equivalent" (see Appendix A for the computation and a description of this variable). Marital status and the number of children are not included in this column. Arguably, this income measure is closer to the individual consumption level. Using this variable or simply household income has very little effect on median household income. Given that both income variables yield similar findings, we will use the latter for the remainder of this research.

Previous results would suggest *a priori* that individuals report lower levels of satisfaction when their neighbors are richer. This remark could be challenged by the fact that our previous work appealed to counties which combine more than one cities. Another issue is that agents living in large cities identify themselves to their neighborhood. Columns 4 to 7 verify if the association between satisfaction and median income depends on the size of the local area which is considered in the analysis.

[10]We also find similar findings when we replace the "Log of Real Median Household Income at the County-Level" by category dummies (i.e. "Less than $30,000", "More than $30,000" but Less than $40,000, ...). Our results suggest that the utility function is decreasing at an increasing rate in (y^*).

First, column 4 restricts the sample to counties having a very large land area (greater than 706 sq mi). In our sample, a bit less than 50% of the respondents live in those counties. 706 sq mi correspond to a radius of 15 miles. Then, column 5 keeps only individuals living in counties having a land area greater than 254 sq mi (radius of 9 miles) but smaller than 706 sq mi (43% of the sample). The sixth column restricts the sample to counties smaller than 254 sq mi but larger than 28.26 sq mi (radius of 3 miles). Approximately 7% of the respondents live in those counties. In column 7, we present an interaction between land area and our variable of interest "Log of Median Real Household Income at the County-Level". We rely on an OLS for this specification since Ai and Norton (2003) pointed out that interpreting interaction terms in a nonlinear model is not straightforward. Nonetheless, using an Ordered Probit yields very similar outcomes. This last specification is our favorite since all the sample is preserved.

We find a negative relationship between well-being and neighbors' income when the scale of analysis is large, but this effect fades away when the size of the area is small. The coefficient of interest is even positive but not statistically significant in column 6. Column 7 shows a large and positive association between satisfaction and land area probably meaning that the quality of life is better in rural areas. The interaction between land area and county median income is negative and statistically significant indicating that the negative impact of relative income is higher in large counties.

In summary, our findings seem to point out that the size of the county is critical in determining the effect of neighbors' income on satisfaction. The next subsection will confirm this intuition by analyzing the association between well-being and ZIP code median income.

B. ZIP Code-Level

We now turn our attention to a different geographical area, ZIP codes (5-digit). Arguably, a ZIP code represents a city or a large neighborhood. Hence, this probably means that residents are more homogeneous than at the county-level. Social distance may also play a critical role. As mentioned before, agents may care more about

residents close to them and less about those in a faraway city.

Our econometrics model is now the following:

$$LS_{ijt} = \alpha + \nu_c + \eta_t + \delta ln(y_{ijt}) + \theta ln(y^*) + \gamma X_{ijt} + \varepsilon_{ijt} \quad (5)$$

where the dependent variable is still life satisfaction for respondent i in year t living in ZIP code j. County and year fixed effects completely control for any fixed differences between counties and between years, which means that only within-county variation is used in the estimation. One caveat of this methodology is that a single ZIP code may span more than one county. To verify that this is not an issue, we also present the findings where state dummies are included instead of county fixed effects. We present robust standard errors clustered at the ZIP code-level.

Table 3 illustrates our findings with ZIP code data. We restrict the sample to ZIP codes where the number of respondents is greater than 50 for statistical reliability. We combine all the years when doing such an exercise. Our variable of interest is the "Log of Median Income at the ZIP Code-Level". Once again, the first two columns present the relationship between satisfaction and median income (ZIP code) with and without demographic characteristics. Strikingly, the association between neighbors' income and well-being is now positive and statistically significant. In other words, respondents living in richer ZIP codes report being more satisfied with their life. This is clearly in contradiction with previous findings in the US literature but in line with the idea that the size of the area matters. The coefficient of interest is almost twice as large in column 1 than in column 2 probably indicating that individual characteristics are correlated with neighborhood characteristics.

The next three columns repeat the exercise of restricting the samples to different geographical areas. Column 3 keeps ZIP codes having a size between 254 and 28.26 sq mi (50% of the sample). The fourth column restricts the sample to ZIP codes having a size smaller than 28.26 and larger than 3.14 sq mi (40% of the sample). Column 5 keeps ZIP codes smaller than 3.14 and larger than 0.94 sq mi (radius of 0.55 miles) which represents only 5% of our sample. Column 6 shows an interaction between the land area and the "Log of Median Income at the ZIP Code-Level". The

findings suggest that the relationship between the median income and satisfaction is nonlinear. The positive effect is larger in column 4 than in columns 3 and 5 and the estimated association is statistically significant only in the fourth column. On the other hand, the interaction in column 6 is positive and insignificant.

We believe the size of the reference group (i.e. ZIP codes or counties) is central for understanding income spillovers since it helps us to disentangle two of the main channels. Our interpretation is that the positive impact of public goods on well-being dominates the negative effect of social status at the ZIP code-level whereas it is the opposite at the county-level. We will verify empirically this theory in the next sub-sections by including the relative rank in the income distribution and public goods in our model.

Table 4 presents the effect of neighbors' income but using the two median incomes (ZIP code and county) simultaneously. Column 1 first replicates the second column of Table 3 for comparison purposes. In column 2, the association between neighbors' income and satisfaction is positive for close neighbors (ZIP code) and negative at the county-level. This is consistent with our previous results. Agents enjoy living in a rich city but also having relatively poorer individuals with whom they do not interact since they have preferences for high social status. Interestingly, the coefficient of the variable "Log of Median Income at the ZIP Code-Level" is twice as small as the coefficient of the variable "Log of Median Income at the County-Level", but both estimates are statistically significant at the 1% level. This indicates that if both county and ZIP-code median incomes rise by the same percentage, a person would be less satisfied. The overall impact of neighbors' income is thus negative when considering both median incomes.

We also verify if neighbors' spillovers have a different effect by age categories. It is plausible that younger agents are more likely to be affected by income comparisons. Column 3 shows two interactions between a dummy indicating if the respondent is younger than 65 years old and the two median incomes. The interactions have different signs and are both statistically significant. While the sign of the interaction between being younger than 65 years old and neighbors' income at the ZIP code-level is positive, this is the opposite at the county-level. The fact that younger agents are

more affected by income spillovers may indicate that amenity and social status are relatively more important for this demographic group.

Finally, the last two columns restrict respectively the sample to respondents having a household income smaller and larger than the median income at the ZIP code-level. We find that both groups enjoy living with rich close neighbors but that the negative association with county median income is larger for poorer respondents. This is an indication that comparisons affect poorer agents relatively more than richer agents but that both benefit from public goods.

C. Rank in the Local Income Distribution

Tables 5 upgrades our basic model by introducing the individual household's normalized rank in the local income distribution (see equations (3) and (4)). Columns 1, 2 and 3 present our estimates at the county-level while columns 4, 5 and 6 do the analysis at the ZIP code-level. Appendix D gives the details about the construction of this variable. The first column of Table 5 reproduces the finding of the second column of Table 2 for comparison purposes. Column 2 includes our measure of social status, the relative rank in the county. Unsurprisingly, the relative rank in the county of residence is positively related to satisfaction. The coefficient is very large and statistically significant at the 1% level. But adding the normalized rank to the model affects dramatically the coefficient on median household income at the county-level. The coefficient is now positive and statistically significant which means that, conditional on own income and the relative rank in the county, respondents report higher levels of life satisfaction when their neighbors are richer. Moreover, adding the relative rank in the local income distribution in a basic well-being regression increases more the pseudo-R-squared than including the median income at the county-level.

This augmented specification complements our previous results. Our interpretation is that once the social status of the respondent is taken into account, richer faraway neighbors are welcome since they provide better amenities. The effect of social status is much larger than the impact of public goods at the county-level and the opposite at the ZIP code-level. Including the relative rank affects also the size

of the coefficient of own household income. The coefficient is divided by four but remains statistically significant. This means that the relative rank captures a fraction of the relative income but also of own income. Agents may care about how many individuals in their reference group are relatively richer and not about by how much they are richer. Comparisons would thus be local in the income distribution. Appendix Table 2 verifies our main findings at the county-level by using different specifications and multiplication factors for the bottom and top income categories. This seems to have little effect on the coefficients of interest.

Table 5, column 4, looks at the effect of income ranking on life satisfaction within a ZIP code. The size of the coefficient of median household income at the ZIP code-level is now larger and remains significant when the rank in the income distribution is included. The relative rank is positive and statistically significant as it was the case with county-level specifications. However, the size of the coefficient is sensitive to sample size (whether we keep only ZIP codes with 50 respondents for a given year) but also to our treatment of top and bottom-coded categories. Appendix Tables 3 and 4 present the same specifications as in columns 4 and 5 of Table 5, but for different multiplication factors for top and bottom categories. The relative rank is positive and statistically significant for 12 out of 16 regressions (not significant when the bottom category is divided by 1.5 and the top category is multiplied by either 2 and 2.5). Nonetheless, the different coefficients of the relative rank are smaller than at the county-level. This is another piece of evidence that at the ZIP code-level, the relative position is far less important than at the county-level.

We then verify whether the rank in the local income distribution matters more at the bottom/top of the distribution. Columns 3 and 6 of Table 5 replace the variable "Normalized Rank" by 5 dummies representing classes of ranking in the county and ZIP code income distribution. We omit the bottom class (having a rank smaller than 0.20) and thus measure the effect of being in a particular class in comparison of being in the bottom class. We find nonlinearities in the relationship between satisfaction and relative ranking. Our results suggest that there is no statistical difference between respondents in the first two classes at the ZIP code-level and a small effect at the county-level. On the other hand, moving from the third to the

fourth class of the county income distribution is predicted to increase life satisfaction 4 times more than moving from the first to the second. This may indicate that being in the bottom of the local income distribution is very detrimental compared to being in the top classes but that this negative effect diminishes only when getting to the middle class. Obviously, any conclusion is tentative since panel data and indication of social mobility are necessary for this type of analysis.

As a pedagogical device, Figures 1, 2 and 3 illustrate the BRFSS county life satisfaction distribution using three different specifications. We estimate an OLS similar to the regression presented in column 4 of Table 2. In this regression, we do not include the rank and the median household income. We then show the average life satisfaction for each county (Figure 1). Figure 2 is similar to Figure 1 with the exception that median household income is included in the specification. Last, Figure 3 adds to the model the rank and the median household income. These figures give the opportunity to have a look at the county well-being distribution in the US, and then to counterfactuals where there would be no income spillover effects.

D. County: Public Goods and Poverty

Since county median income is negatively related to utility, then controlling for public goods should decrease the size of the coefficient of y^* as explained in Section II. It is also possible that local omitted variables such as values of housing units explain our findings that neighbors' income affect well-being. This would mean that the relationship between well-being and income spillovers is spurious. If this is the case, then including county-level variables in our model would make the coefficient of interest insignificant. Table 6 verifies the validity of our previous findings by introducing county-level variables in the specification (see Appendix B for the definitions of these variables).

Column 1 introduces the county unemployment rate, the owner-occupied housing units and the median value of specified owner-occupied housing units to capture any business cycle effects. It is possible that the intensity of comparisons is larger/smaller during recessions.[11] Moreover, including the median value of speci-

[11] We do not find any evidence that income comparisons are more important before or after the

fied owner-occupied housing units and other county-level variables allows us to tackle the omitted variable bias. The second column includes the percentage of elderly and the percentage of high school graduate or higher (25 years and older). Poterba (1997) found that a larger fraction of elderly in a jurisdiction leads to lower public spending on education. It is also possible that networking with people with higher education raises well-being. Column 3 adds to the basic model the density of population and the urbanization rate since it is possible that social interactions are more prevalent in small cities. Column 4 includes the number of murders and nonnegligent manslaughters known to police per capita (using robberies or violent crimes instead yields similar findings). Column 5 includes all these controls in the model. Columns 6, 7 and 8 includes other types of county-level characteristics which are more likely to be perceived as public goods: local government direct expenditures on health, education, welfare, total expenditures and local revenue.

Adding these controls in the model does affect the size of the coefficient of interest significantly. Agents report lower levels of satisfaction when their neighbors are richer, but this association is larger when county-level characteristics are included. Controlling for amenities in the local area decreases the size of the effect of neighbors' income on agents' utility (especially when the full set of county-level variables is included, column 5). This is another piece of evidence that mainly two effects affect the coefficient of y^*: social status and amenities.

Interestingly, characteristics such as criminality and the economic situation seem to affect more the variable "Log of Median Real Household Income at the County-Level" than direct expenditures by the local government. This might be an indication that agents prefer to live with rich neighbors mainly because it provides a safer environment and better economic conditions.

Moreover, public goods explain why conditional on own income and the relative rank, agents report higher levels of utility when their neighbors are richer. Including the public goods in our model impacts only the coefficient of median income and

Great Recession. We verified this hypothesis by restricting the sample to the period 2005-2007 in one regression and 2008-2010 in another (not shown). The coefficients for the variable "Log of Median Real Household Income at the County-Level" are quite similar and thus not statistically different from each other.

not the coefficient of the relative rank in the income distribution (not shown).

Our findings do not exclude the possibility that agents' utility is decreasing both with the prevalence of rich and poor households. This would mean that individuals are sensitive to poverty and prefer living in counties where they are among the richest people, but also where poverty is not prevalent. Table 7 looks at the relationship between satisfaction and different measures of inequality and poverty at the county-level. Columns 1, 2, and 3 enrich equation (2) by adding respectively the percentage of people of all ages in poverty, the percentage of related children age 5 to 17 in families in poverty and the percentage of people under age 18 in poverty in the county. Conditional on own income and neighborhood median income, individuals report lower levels of satisfaction as the proportion of poor people increases. Columns 4 and 5 restrict respectively the sample to counties having a Gini coefficient below and above the median (0.441). The effect of median income is similar for both columns and not significantly different. Note that the first five columns do not include any county-level variables.

Columns 6, 7 8, 9 and 10 add to the previous specifications the number of murders and nonnegligent manslaughters known to police per capita (we also add other county-level variables in some specifications, available upon request). There is no more clear evidence that respondents are averse to poverty when we include amenities. It seems that this aversion to poverty is driven by aversion to criminality and other factors associated with poverty.

IV. Somerville

A. Somerville Phone Survey

While the previous section analyzed interactions within counties and ZIP codes, this section looks at a more finely disaggregated level: streets. It may be the case that jealousy has a stronger effect with people with whom we interact very often such as street neighbors. In other words, we analyze whether respondents' well-being is affected by standard of living of individuals living in the same street.

This part of our analysis focuses on a survey designed by the City of Somerville's SomerStats Office[12]. Somerville became the first city in the US to collect data about residents' self-reported well-being. While this city is very particular and not representative of the US, we believe using such a disaggregated dataset will give a first piece of evidence of the presence of income comparisons at the street-level. Previous empirical papers in this literature have been relying on very large areas as a reference group and thus probably captured the effects of many other channels.

Somerville is a city in Middlesex County, Massachusetts, located just north of Boston. As of the 2010 census, the city had a total population of 75,754 and was the most densely populated municipality in New England with 18,404 people per square mile. The median income for a household in the city was $61,731 for the period 2006-2010. The largest industry sectors in terms of employment are health and social services. Approximately 17% of Somerville residents work within the city while 50% work in Boston or Cambridge.[13]

Somerville Phone Survey ("SMP" hereafter) has been conducted during the spring of 2011 via phone, email and Facebook.[14] A total of 393 respondents answered questions such as "How satisfied are you with your life in general?" and "How happy do you feel right now?" where respondents have 10 choices. Many respondents did not answer key questions such as household income, life satisfaction or address information which leaves us with 323 individuals. Table 1 shows means and standard deviations of these variables. 14% and 10.8% of the respondents report respectively that they are very satisfied with their life and very happy. On the other hand, less than 1% are very dissatisfied or unhappy. The household income question has 16 choices going from less than $20,000 to $300,000 and more. Respectively, 11% and

[12]The data from this survey are derived from Sensitive Data Files of the City of Somerville, obtained under special arrangements designed to protect the anonymity of respondents. These data are not available from the authors. Persons interested in obtaining Data files should contact the City of Somerville. Contact information is available on our websites.

[13]See this website for technical reports on the economic situation: http://www.somervillema.gov/spotlights/comp-plan/trends-meetings-and-reports.

[14]See Heller (2011) for more details on the SMP. The questionnaire is available on our websites. Another survey, the Somerville's Well-Being and Community Survey, has been conducted during the same period. This survey is a special well-being questionnaire to households as part of the annual city census. It has a total sample of 6,167 respondents and included questions on household income and self-reported well-being. Unfortunately, the address of residence is not available which explains our choice to focus on the SMP.

1% of the individuals report having less than $20,000 and more than $300,000. We use multiplication factors as before in order to have an income distribution closer to real figures. This survey also contains information on gender, age, race, marital status, years living in Somerville, ward of residence (7 choices), precinct (21 choices), ZIP code (3 choices)[15] and employment status (see Table 1).

The household median income of the respondents in the SMP is $70,000. This is higher than the the median household income provided in the Census which could indicate that the survey is not representative. Appendix Table 5 compares the socioeconomic characteristics of the respondents to those of the 2010 Census and the Community Survey. There is an under-representation of the poor families and an over-representation of women and persons 65 years old and over. Readers should thus keep in mind that the findings of this section have less external validity. Yet, using such a survey is helpful in understanding income comparisons at a very disaggregated level.

B. Street-Level Analysis

We estimate social comparisons in the City of Somerville by measuring the distance between the 323 residents.[16] The survey provides information about respondents' place of residence which allows us to measure the exact distance between the 323 respondents, using longitude and latitude coordinates. We match the address of residence with different measures of public goods. We calculate the distances between respondents' residence and different local amenities such as subway stations, parks and libraries (Heller (2011)). Appendix F describes these variables and provides the different sources where we obtained the geolocalisation.

Our strategy involves matching a group of neighbors for each citizen. This group

[15] www.city-data.com/zipmaps/Somerville-Massachusetts.html.

[16] We also adopt an estimation strategy that enables us to restrict the comparison group to individuals living in the same ward (not shown for space consideration). Somerville is composed of 7 wards of similar population magnitude which allows us to look at whether individuals report higher or lower satisfaction levels when their ward neighbors are richer/poorer. In this matching framework, our standard errors are clustered at the ward-level. The estimated effects of the reference income (i.e ward median household income) on life satisfaction is negative and significant. This means that conditional on own income, respondents report lower levels of satisfaction when their ward neighbors are richer (not shown).

is composed of residents within a circle of radius r. A first circle of 0.20 miles radius (0.12 sq mi) is centered around each respondent. This circle contains approximately 10% of the sample. We also construct another circle of 0.40 miles radius, but we keep only respondents living between the two distances (0.20 and 0.40 mi) in order to build our second reference group. This ring contains 10% of the sample. Figure 4 illustrates our empirical strategy.

Our econometrics model is:

$$SWB_{ir} = \alpha + \delta ln(y_i) + \theta ln(y_r^*) + \gamma X_i + \varepsilon_{ir} \tag{6}$$

where life satisfaction is the dependent variable for individual i in a given circle having a radius r, δ is the coefficient associated with household income, X is a set of individual covariates and θ is the coefficient of interest. We present robust standard errors clustered at the precinct-level. Individual covariates include age dummies, gender, 4 dummies of working status (employed, unemployed, student and out of the labor force), 5 dummies of marital status (married, single, living as couple, divorced, separated or widowed) and 4 race dummies (White, Asian, American Indian and African American).

Table 8 presents our basic findings of equation (6). In the first 3 columns, life satisfaction is the dependent variable whereas it is happiness in the last three columns. Columns 1 and 4 show the relationship between self-reported well-being and median household income for the first circle only. The coefficient of our variable of interest "Log of Median Household income (0 to 0.12 sq mi)" is negative in both happiness and satisfaction regressions but statistically significant solely in the former. This means that, conditional on own income, respondents report lower levels of happiness when their street neighbors are richer. The size of the coefficient is larger than at the county-level and is even bigger than the coefficient of own household income in the happiness regression.

Columns 2 and 5 upgrade our basic model by including the ring (0.2 to 0.4 mi) instead of the first circle. Recall that only the income of the respondents between the distances is included. The coefficient of the second reference group is positive

but not statistically significant in both regressions. This means that respondents compare mainly to next door neighbors and less to other citizens of Somerville. These findings are not sensitive to the use of a multiplication factors for top/bottom income categories (see Appendix Table 6). Columns 3 and 5 include at the same time the circle and the ring which yields somewhat similar findings.

We control for distances to the nearest park, public library and subway station in Table 9.[17] We obtained these measures for each survey point from geocoding data (Heller (2011)). We also include an index of housing price in the regressions. We use the median price of recent home sales near the place of residence. The data come from public records publicly available on HomeInsight's website (see Appendix F for more details). Residents of Somerville have access to this data easily on HomeInsight.com and are thus potentially aware of the median price we are using.

The upper panel of Table 9 shows the results using life satisfaction while the bottom relies on happiness as the dependent variable. Column 1 presents the specification including the distance to the nearest subway whereas columns 2 and 3 include respectively the distances to the nearest park and library. Last, column 4 includes the three distances to amenities. Median housing prices are included in all the columns. Overall, controlling for distances to amenities in the local area and median price of recent sales does not affect our previous conclusion that next door neighbors' income is negatively associated to respondents' happiness.

The findings of Somerville point out the fact that income comparison is present and affects agents' utility. We believe income comparison has a stronger effect on well-being at the street-level than at the county and ZIP code-level since agents observe directly the consumption of their next door neighbors. Nevertheless, this survey has a small sample size and is not representative. Any conclusions must be tentative. We believe this section provided a first step towards identifying the impact of income comparison on well-being using street-level data.

[17] Sirgy and Cornwell (2002) argue that satisfaction with physical features affects both neighborhood satisfaction and housing satisfaction. Satisfaction with the social features of the neighborhood plays a role in the satisfaction with the neighborhood which affects life satisfaction.

V. Conclusion

This paper analyses the relationship between well-being and neighbors' income. Our main findings suggest that the size of the area affects drastically this relationship. In smaller areas such as ZIP codes, individuals report higher levels of well-being when their neighbors are rich whereas it is the opposite at the county-level. This is consistent with the idea that agents enjoy living in a rich city but also having relatively poorer faraway individuals since they have preferences for high social status. We test explicitly our interpretation by including amenities and the relative rank in the income distribution in our model.

We also find evidence that residents of Somerville compare to people living in their street. Hence, the relationship between well-being and income spillovers follows an inverted U-shaped pattern in the size of the area. Table 10 gives a summary of our main findings.

While the American literature has concentrated on the psychological role of relative consumption, we have shown that amenities matter in determining the effect of neighbors' income on utility. Obviously, the role of relative consumption is also important and is very strong for next door neighbors. These results point to some interesting future research questions. Due to a lack of longitudinal data over a very long period of time, it was not possible to test whether individuals who experienced social mobility are more prone to comparisons. It is also of general interest to know if policies aiming at reducing income inequality are improving general welfare.

References

Ai, C. and Norton, E. C.: 2003, Interaction terms in logit and probit models, *Economics Letters* **80**(1), 123–129.

Angrist, J. D. and Pischke, J.-S.: 2009, Mostly harmless econometrics: An empiricist's companion, *Princeton University Press* .

Barrington-Leigh, C. P. and Helliwell, J. F.: 2008, Empathy and emulation: Life satisfaction and the urban geography of comparison groups, *NBER Working Papers* (14593).

Blanchflower, D. G. and Oswald, A. J.: 2004, Well-being over time in britain and the usa, *Journal of Public Economics* **88**(7-8), 1359–1386.

Boskin, M. J. and Sheshinski, E.: 1978, Optimal redistributive taxation when individual welfare depends upon relative income, *Quaterly Journal of Economics* **92**, 589–601.

Brodeur, A.: 2012, Smoking, income and subjective well-being: Evidence from smoking bans, *PSE Working Papers, HAL* (halshs-00664269).

Brown, D. A., Gardner, J., Oswald, A. and Quian, J.: 2008, Does wage rank affect employees' well-being?, *Industrial Relations* **47**(3), 355–389.

Card, D., Mas, A., Moretti, E. and Saez, E.: 2012, Inequality at work: The effect of peer salaries on job satisfaction, *American Economic Review* **102**(6), 2981–3003.

Clark, A. E., Frijters, P. and Shields, M. A.: 2008, Relative income, happiness, and utility: An explanation for the easterlin paradox and other puzzles, *Journal of Economic Literature* **46**(1), 95–144.

Clark, A. E., Kristensen, N. and Westergard-Nielsen, N.: 2009, Economic satisfaction and income rank in small neighbourhoods, *Journal of the European Economic Association* **7**(2-3), 519–527.

Clark, A. E. and Oswald, A. J.: 1996, Satisfaction and comparison income, *Journal of Public Economics* **61**(3), 359–381.

Clark, A. E. and Senik, C.: 2010, Who Compares to Whom? The Anatomy of Income Comparisons in Europe, *Economic Journal* **120**(544), 573–594.

DiTella, R. and MacCulloch, R.: 2006, Some uses of happiness data in economics, *Journal of Economic Perspectives* **20**(1), 25–46.

Dittmann, J. and Goebel, J.: 2010, Your house, your car, your education: The socioeconomic situation of the neighborhood and its impact on life satisfaction in germany, *Social Indicators Research* **96**(3), 497–513.

Duesenberry, J. S.: 1949, Income, saving and the theory of consumer behavior, *Harvard University Press, Cambridge, MA.* .

Easterlin, R. A.: 1995, Will raising the incomes of all increase the happiness of all?, *Journal of Economic Behavior and Organization* **27**, 35–48.

Frank, R. H.: 1985, The demand for unobservable and other nonpositional goods, *American Economic Review* **75**(1), 101–16.

Frey, B. S. and Stutzer, A.: 2002, What can economists learn from happiness research?, *Journal of Economic Literature* **40**(2), 402–435.

Goudie, R. J. B., Mukherjee, S., DeNeve, J.-E., Oswald, A. J. and Wu, S.: 2011, Happiness as a driver of risk-avoiding behavior, *CESifo Working Paper Series 3451*, CESifo Group Munich.

Grosfeld, I. and Senik, C.: 2010, The emerging aversion to inequality, *The Economics of Transition* **18**(1), 1–26.

Heller, N.: 2011, Resident happiness and urban amenities in somerville, ma: A geographic analysis, *ISMT E-150 Geographic Information Systems Final Project Report* .

Helliwell, J. F. and Huang, H.: 2011, New measures of the costs of unemployment: Evidence from the subjective well-being of 2.3 million americans, *National Bureau of Economic Research* (16829).

Hirschman, A. O. and Rothschild, M.: 1973, The Changing Tolerance for Income Inequality in the Course of Economic Development; with a Mathematical Appendix, *The Quarterly Journal of Economics* **87**(4), 544–66.

Kapteyn, A., de Geer, S. V. and de Stadt, H. V.: 1985, The relative utility; Evidence from panel data, *Review of Economics and Statistics* **67**, 179–187.

Kingdon, G. G. and Knight, J.: 2007, Community, comparisons and subjective well-being in a divided society, *Journal of Economic Behavior & Organization* **64**, 69–90.

Knies, G., Burgess, S. and Propper, C.: 2008, Keeping up with the schmidt's: An empirical test of relative deprivation theory in the neighbourhood context, *Schmollers Jahrbuch : Journal of Applied Social Science Studies* **128**(1), 75–108.

Knight, J., Song, L. and Gunatilake, R.: 2009, Subjective well-being and its determinants in rural china, *China Economic Review* **20**(4), 635–649.

Kuhn, P., Kooreman, P., Soetevent, A. and Kapteyn, A.: 2011, The Effects of Lottery Prizes on Winners and Their Neighbors: Evidence from the Dutch Postcode Lottery, *American Economic Review* **101**(5), 2226–47.

Luttmer, E. F. P.: 2005, Neighbors as negatives: Relative earnings and well-being, *The Quarterly Journal of Economics* **120**(3), 963–1002.

Oswald, A. J. and Wu, S.: 2010, Objective confirmation of subjective measures of human well-being: Evidence from the usa, *Science* **327**.

Pfeffermann, D.: 1993, The role of sampling weights when modeling survey data, *International Statistical Review, International Statistical Institute* **61**(2), 317–337.

Poterba, J. M.: 1997, Demographic structure and the political economy of public education, *Journal of Policy Analysis and Management* **16**(1), 48–66.

Powdthavee, N.: 2009, How important is rank to individual perception of economic standing? A within-community analysis, *Journal of Economic Inequality* **7**(3), 225–248.

Robson, A.: 1992, Status, the distribution of wealth, private and social attitudes to risk, *Econometrica* **60**, 837–57.

Sirgy, M. J. and Cornwell, T.: 2002, How neighborhood features affect quality of life, *Social Indicators Research* **59**, 79–114.

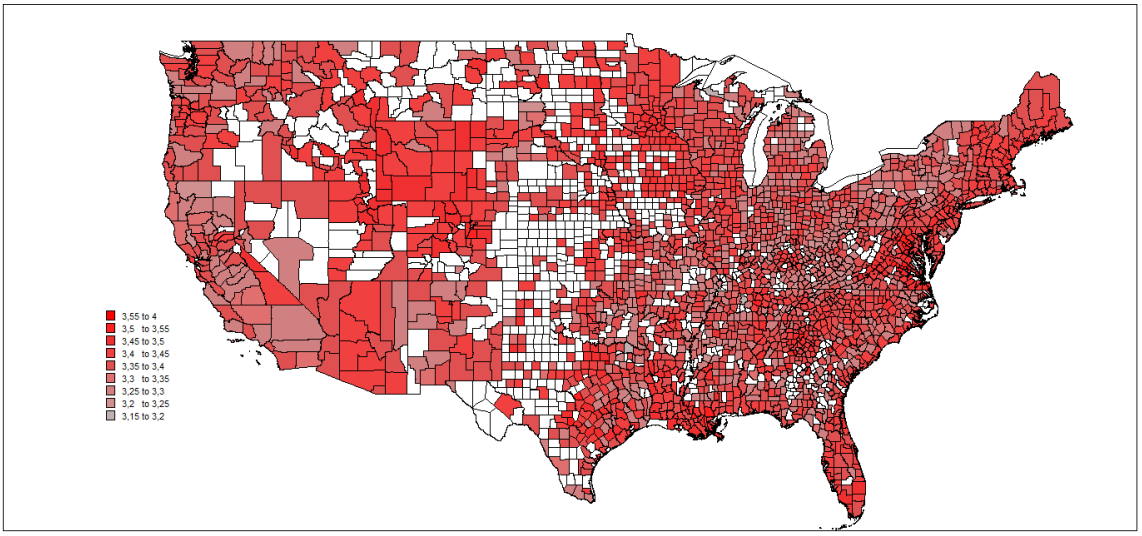
Stigler, G.: 1950, The development of utility theory, i and ii, *Journal of Political Economy* **58**, 373.

VanPraag, B. M. S.: 1971, The welfare function of income in Belgium: An empirical investigation, *European Economic Review* **2**, 337–369.

Veblen, T. B.: 1899, The theory of the leisure class, *London: Macmillanm George Allen and Unwin* .

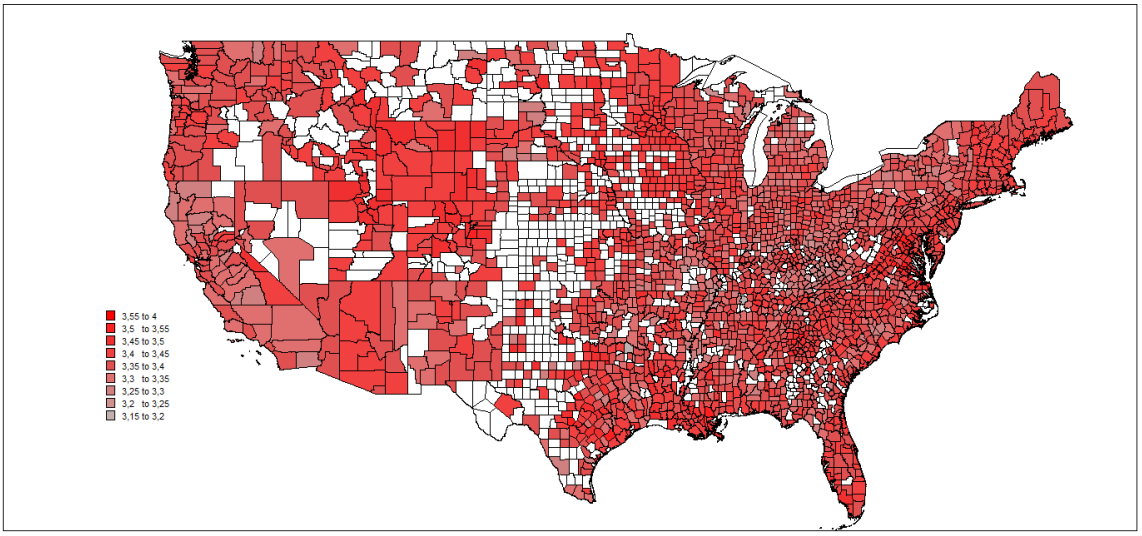
VI. Figures

Figure 1: BRFSS County Life Satisfaction Distribution



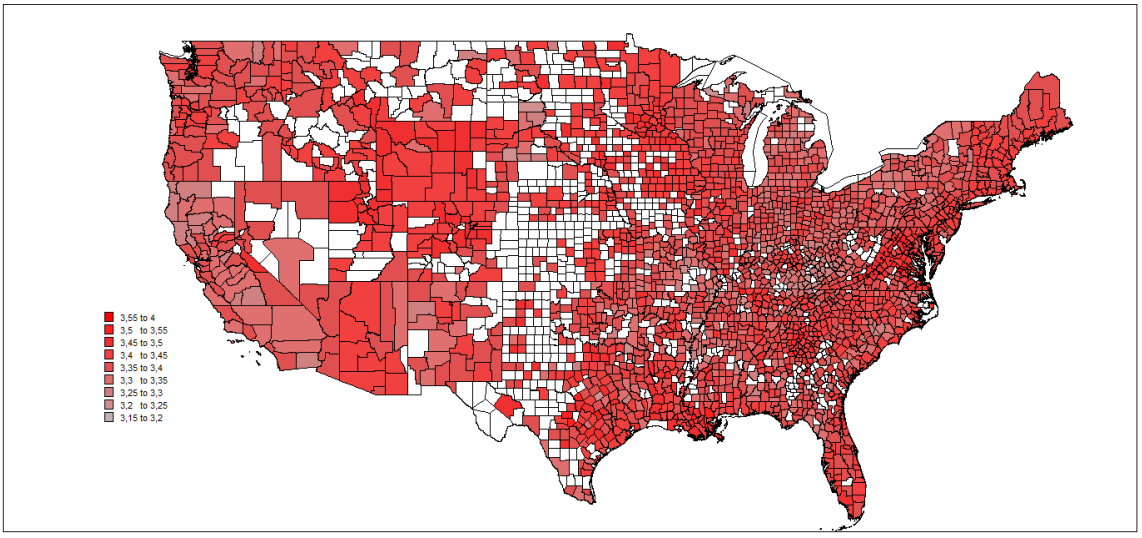
Notes: Data came from the BRFSS, 2005-2010. Blank means there were no data for this county. This figure illustrates the BRFSS county life satisfaction distribution. We estimate an OLS similar to the regression presented in Table 2 with the exception that we do not include the median household income. We then present the average life satisfaction for each county.

Figure 2: BRFSS County Life Satisfaction Distribution



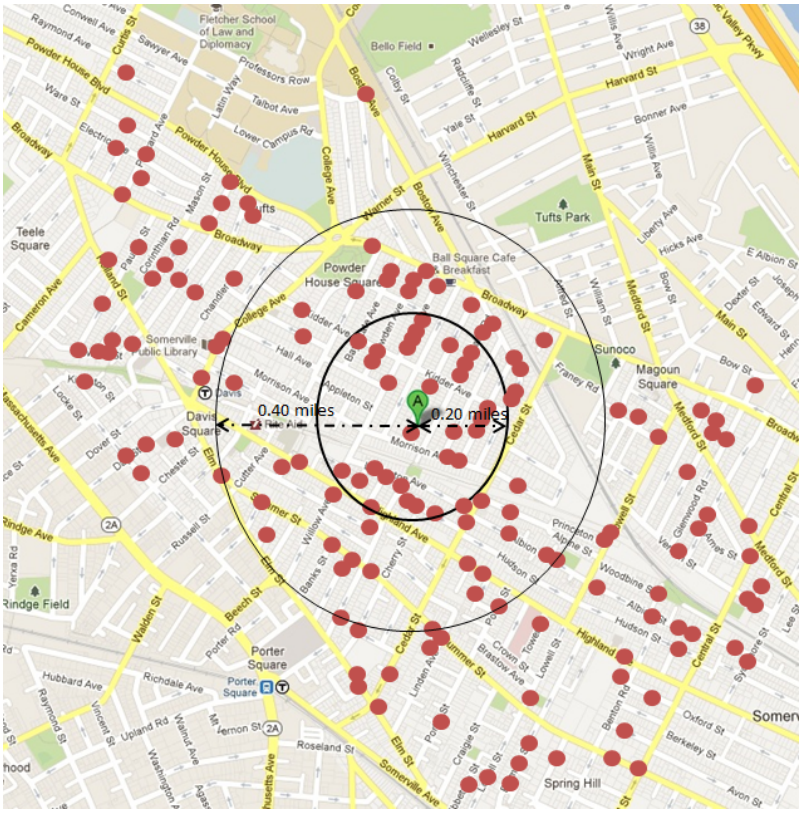
Notes: Data came from the BRFSS, 2005-2010. Blank means there were no data for this county. This figure illustrates the BRFSS county life satisfaction distribution where income spillovers are taken into account. We estimate an OLS similar to the regression presented in column 2 of Table 2. In this regression, we do include the median household income. We then present the average life satisfaction for each county.

Figure 3: BRFSS County Life Satisfaction Distribution



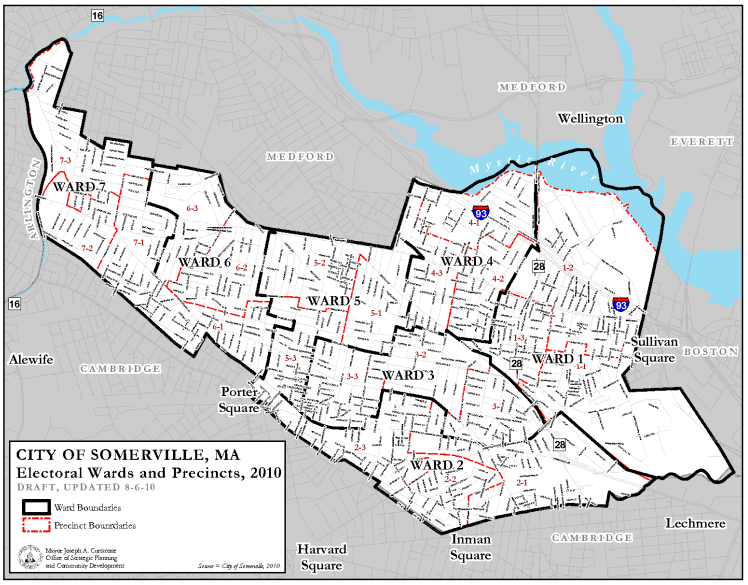
Notes: Data came from the BRFSS, 2005-2010. Blank means there were no data for this county. This figure illustrates the BRFSS county life satisfaction distribution where income spillovers and ranking in the income distribution are taken into account. We estimate an OLS similar to the regression presented in column 2 of Table 4. In this regression, we do include the rank and the median household income. We then present the average life satisfaction for each county.

Figure 4: Somerville Phone Survey



Notes: Data came from the City of Somerville (not available from the authors). We distinguish a specific group of reference for each respondent: neighbors within a circle of radius r. A first circle of 0.20 miles radius is centered around each respondent. We then construct another circle of 0.40 miles radius, but we keep only respondents living between the two distances (i.e. ring).

Figure 5: City of Somerville: Electoral Wards and Precincts



Notes: This Figure illustrates the geolocalisation of the 21 precincts in Somerville. There are 21 precincts in Somerville (3 per ward). Each precinct contains on average 5% of the respondents. This map and many others are available on the City of Somerville's website: http://faqs.somervillema.intelligovsoftware.com/wardandprecintmaponline.aspx.

VII. Tables

Table 1: Summary Statistics.

BRFSS	Mean	Std. Dev.	**SMP**	Mean	Std. Dev.
Life Satisfaction			Life Satisfaction		
[1] Very Dissatisfied	3.389	0.622	[1] Very Dissatisfied	7.687	1.847
[4] Very Satisfied			[10] Very Satisfied		
Very Satisfied	0.453	0.498	Very Satisfied	0.139	0.346
Very Dissatisfied	0.010	0.101	Very Dissatisfied	0.006	0.078
			Happiness		
			[1] Very Unhappy	7.470	1.835
			[10] Very Happy		
			Very Happy	0.108	0.311
			Very Unhappy	0.006	0.078
Household Income			Household Income		
[0,10 000[	0.048	0.214	[0,20 000[	0.111	0.315
[10 000,15 000[	0.049	0.215	[20 000,40 000[	0.158	0.365
[15 000,20 000[	0.068	0.252	[40 000,60 000[	0.130	0.336
[20 000,25 000[	0.085	0.278	[60 000,80 000[	0.158	0.365
[25 000,35 000[	0.111	0.315	[80 000,100 000[	0.130	0.336
[35 000,50 000[	0.148	0.355	[100 000,120 000[	0.102	0.303
[50 000,75 000[	0.171	0.376	[120 000,140 000[	0.062	0.241
75 and more	0.321	0.467	[140 000,160 000[	0.037	0.189
Male	0.498	0.500	[160 000,180 000[	0.031	0.173
Age	45.99	16.861	[180 000,200 000[	0.018	0.135
Elementary School	0.035	0.184	[200 000,220 000[	0.018	0.035
Att. High School	0.064	0.245	[220 000,240 000[	0.006	0.078
Grad. High School	0.270	0.444	[240 000,260 000[	0.009	0.096
Att. College	0.268	0.443	[260 000,280 000[	0.006	0.078
Grad. College	0.364	0.481	[280 000,300 000[	0.009	0.096
Married	0.618	0.486	300 000 and more	0.012	0.087
Divorced	0.091	0.288	Female	0.582	0.493
Single	0.292	0.379	Couple	0.114	0.318
Separated	0.174	0.142	Married	0.452	0.498
Widowed	0.056	0.231	Divorced or Separated	0.108	0.311
Couple	0.040	0.196	Widowed	0.068	0.252
No Child	0.557	0.497	Employed	0.614	0.487
One Child	0.173	0.379	Unemployed	0.061	0.241
Two Children	0.167	0.373	Out of Labor Force	0.226	0.418
Three Children or more	0.103	0.304	Student	0.040	0.196
Employed	0.541	0.498	White	0.818	0.386
Unemployed (Less 1Y)	0.036	0.186	African American	0.034	0.181
Unemployed (More 1Y)	0.024	0.154	Asian/Pacific Islander	0.018	0.135
Self-Employment	0.088	0.283	American Indian	0.003	0.055
Retired	0.149	0.356			
Disabled	0.048	0.213			
Student	0.041	0.197			
Full-Time Homemaker	0.074	0.261			
White	0.772	0.419			
Black or African American	0.106	0.308			
Asian	0.034	0.180			
Pacific Islander (Hawaiian)	0.004	0.065			
American Indian or Alaska	0.015	0.123			
Other Race	0.041	0.199			
Multiracial	0.020	0.141			
Observations	1,737,499		Observations	323	

Note: For the BRFSS, sample means are weighted using the final weight associated to each respondent. The period covered is 2005-2010. For the SMP, we do not report means when respondents refused to answer.

Table 2: Life Satisfaction and Income Spillovers at the County-Level, BRFSS.

Life Satisfaction	Ordered Probit (1) All	Ordered Probit (2) All	Ordered Probit (3) All	Ordered Probit (4) .>706	Ordered Probit (5) 706>.>254	Ordered Probit (6) 254>.>28.26	OLS (7) All
Ln (Real HH Income)	0.367	0.236		0.221	0.268	0.214	0.122
	(0.009)	(0.007)		(0.010)	(0.004)	(0.015)	(0.004)
Ln (Real Equiv. HH Income)			0.231				
			(0.007)				
Ln (Median County HH Income)	-0.046	-0.067	-0.060	-0.103	-0.083	0.052	-0.026
	(0.018)	(0.015)	(0.015)	(0.017)	(0.017)	(0.039)	(0.008)
Land Area County (10^3 sqm)							0.095
							(0.038)
Ln (Median County HH Income)* Land area							-0.009
							(0.003)
Control Variables							
Socioeconomic Controls		✓	✓	✓	✓	✓	✓
State Dummies	✓	✓	✓	✓	✓	✓	✓
Year Dummies	✓	✓	✓	✓	✓	✓	✓
Observations	1,737,499	1,737,499	1,737,499	855,957	744,329	131,288	1,737,499
Pseudo R2	0.0380	0.0641	0.0535	0.0638	0.0686	0.0610	0.1142

Note: All estimates are weighted using the final weight associated to each respondent. Robust standard errors are in parentheses, clustered by county. The period covered is 2005-2010. In the first column, only state and year dummies are included. The remaining columns include socioeconomics variables (described in the text). Household income has 8 categories. The log of the real household income is calculated using the middle point of each category (see Appendix).

Table 3: Life Satisfaction and Income Spillovers at the ZIP Code-Level, BRFSS.

Life Satisfaction	Ordered Probit (1) All	Ordered Probit (2) All	Ordered Probit (3) 254> .>28.26	Ordered Probit (4) 28.26>.>3.14	Ordered Probit (5) 3.14>.>0.94	OLS (6) All
Ln (Real HH Income)	0.402	0.272	0.268	0.285	0.240	0.137
	(0.004)	(0.005)	(0.007)	(0.008)	(0.039)	(0.002)
Ln (Median ZIP Code HH Income)	0.121	0.062	0.037	0.078	0.048	0.026
	(0.015)	(0.012)	(0.031)	(0.019)	(0.066)	(0.007)
Land Area ZIP Code (10^2 sqm)						-0.087
						(0.112)
Ln (Median ZIP Code HH Income)* Land area						0.009
						(0.010)
Control Variables						
Socioeconomic Controls		✓	✓	✓	✓	✓
County Dummies	✓	✓	✓	✓	✓	✓
Year Dummies	✓	✓	✓	✓	✓	✓
Observations	221,110	221,110	103,154	79,731	4,735	221,110
Pseudo R2	0.0475	0.0735	0.0736	0.0775	0.0824	0.1302

Note: Robust standard errors are in parentheses, clustered by ZIP Code. The period covered is 2005-2010 (except for Texas (2007-2010)). In the first column, only county and year dummies are included. In the remaining colums, state and year dummies are included in the model in addition to socioeconomics variables (described in the text). Household income has 8 categories. The log of the real household income is calculated using the middle point of each category (see Appendix)

Table 4: Life Satisfaction and Income Spillovers by Socioeconomic Characteristics, BRFSS.

Life Satisfaction	Ordered Probit (1) All	Ordered Probit (2) All	OLS (3) Age	Ordered Probit (4) Poorer	Ordered Probit (5) Richer
Ln (Real HH Income)	0.272	0.271	0.135	0.202	0.345
	(0.005)	(0.005)	(0.002)	(0.008)	(0.012)
Ln (Median ZIP Code HH Income)	0.062	0.051	-0.005	0.063	0.062
	(0.015)	(0.016)	(0.009)	(0.018)	(0.023)
Ln (Median County HH Income)		-0.110	-0.015	-0.144	-0.086
		(0.022)	(0.016)	(0.027)	(0.031)
Ln (Median ZIP Code HH Income)*Age<65			0.039		
			(0.012)		
Ln (Median County HH Income) *Age<65			-0.057		
			(0.019)		
Age<65			0.099		
			(0.160)		
Control Variables					
Socioeconomic Controls	✓	✓	✓	✓	✓
County Dummies	✓				
State Dummies		✓	✓	✓	✓
Year Dummies	✓	✓	✓	✓	✓
Observations	221,110	219,144	219,144	105,190	113,952
Pseudo R2	0.0735	0.0721	0.1277	0.0570	0.0393

Note: Robust standard errors are in parentheses, clustered by ZIP code. The period covered is 2005-2010 (except for Texas (2007-2010)). State and year dummies are included in the model in addition to socioeconomics variables (described in the text). Household income has 8 categories. The log of the real household income is calculated using the middle point of each category (see Appendix). The last two columns restrict respectively the sample to respondents having a household income smaller and larger than the median income at the ZIP code-level.

Table 5: Life Satisfaction, Income Spillovers and Rank, BRFSS.

Life Satisfaction	Ordered Probit (1) All	Ordered Probit (2) All	Ordered Probit (3) All	Ordered Probit (4) All	Ordered Probit (5) All	Ordered Probit (6) Alll
Ln (Real HH Income)	0.236	0.064	0.146	0.272	0.231	0.239
	(0.007)	(0.014)	(0.010)	(0.005)	(0.010)	(0.009)
Ln (Median	-0.067	0.156	0.073			
County HH Income)	(0.015)	(0.022)	(0.018)			
Ln (Median				0.062	0.103	0.099
ZIP Code HH Income)				(0.015)	(0.023)	(0.021)
Relative Rank in County (Col. 2)		0.627			0.143	
ZIP Code (Col. 5)		(0.036)			(0.032)	
Relative Rank:						
Less than 0.20			Omitted			Omitted
[0.20, 0.40[			0.025			-0.002
			(0.008)			(0.009)
[0.40, 0.60[			0.102			0.038
			(0.011)			(0.013)
[0.60, 0.80[			0.217			0.075
			(0.014)			(0.017)
More than 0.80			0.280			0.090
			(0.021)			(0.021)
Control Variables						
Socioeconomic Controls	✓	✓	✓	✓	✓	✓
State Dummies	✓	✓	✓			
County Dummies				✓	✓	✓
Year Dummies	✓	✓	✓	✓	✓	✓
Observations	1,737,499	1,737,499	1,737,499	221,110	221,110	221,110
Pseudo R2	0.0641	0.0650	0.0650	0.0735	0.0736	0.0736

Note: For columns 1, 2 and 3, estimates are weighted using the final weight associated to each respondent. Robust standard errors are in parentheses, clustered by county. The period covered is 2005-2010. State and year dummies are included in the model in addition to socioeconomics variables (described in the text). Household income has 8 categories. The log of the real household income is calculated using the middle point of each category (see Appendix). For columns 4, 5 and 6, robust standard errors are in parentheses, clustered by ZIP Code. The period covered is 2005-2010 (except for Texas (2007-2010)). County and year dummies are included in the model in addition to socioeconomics variables (described in the text). Household income has 8 categories. The log of the real household income is calculated using the middle point of each category (see Appendix).

Table 6: Life Satisfaction, Income Spillovers and Public Goods, BRFSS.

Life satisfaction	Ordered Probit (1)	Ordered Probit (2)	Ordered Probit (3)	Ordered Probit (4)	Ordered Probit (5)	Ordered Probit (6)	Ordered Probit (7)	Ordered Probit (8)
Ln (Real HH Income)	0.237	0.237	0.236	0.237	0.237	0.237	0.236	0.237
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Ln (Median	-0.098	-0.079	-0.070	-0.087	-0.131	-0.085	-0.071	-0.076
County HH Income)	(0.020)	(0.019)	(0.014)	(0.014)	(0.028)	(0.014)	(0.014)	(0.014)
Control Variables								
Unemp. Rate and Owner-Occupied Housing	✓				✓			
Median prices of housing units	✓				✓			
% of Elderly and % High School Graduate		✓			✓			
Population Density and Urban Population			✓		✓			
Criminality				✓	✓			
Local Gvt. Direct Exp.: Health, Education and Welfare						✓		✓
Local Gvt. Revenue and Direct Expenditure							✓	✓
Socioeconomic Controls	✓	✓	✓	✓	✓	✓	✓	✓
State Dummies	✓	✓	✓	✓	✓	✓	✓	✓
Year Dummies	✓	✓	✓	✓	✓	✓	✓	✓
Observations	1,737,499	1,737,499	1,737,499	1,737,499	1,737,499	1,737,499	1,737,499	1,737,499
Pseudo R2	0.0642	0.0641	0.0641	0.0641	0.0642	0.0641	0.0641	0.0641

Note: All estimates are weighted using the final weight associated to each respondent. Robust standard errors are in parentheses, clustered by county. The period covered is 2005-2010. All the columns include state and year dummies in addition to socioeconomics variables (described in the text). Column one includes the unemployment rate at the county-level, the number of owner-occupied housing units per inhabitant and the median value of specified owner-occupied housing units. The second column includes the percentage of population 65 years and over and the percentage of high school graduate. The third column includes the urbanization rate and the population density. Column four includes one index of criminality: the number of murders and nonnegligent manslaughters known to police per capita. Column 6 includes three measures of local government direct expenditures per capita. Column 7 includes local government general revenue and direct expenditures per capita. See Appendix for more details about these county-level variables.

Table 7: Life Satisfaction, Income Spillovers and Poverty, BRFSS.

Life Satisfaction	Ordered Probit (1) All	Ordered Probit (2) All	Ordered Probit (3) All	Ordered Probit (4) Gini > 0.441	Ordered Probit (5) Gini<0.441	Ordered Probit (6) All	Ordered Probit (7) All	Ordered Probit (8) All	Ordered Probit (9) Gini > 0.441	Ordered Probit (10) Gini<0.441
Ln (Real HH Income)	0.237	0.237	0.237	0.230	0.250	0.237	0.237	0.237	0.230	0.251
	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)
Ln (Median	-0.118	-0.141	-0.140	-0.072	-0.074	-0.106	-0.124	-0.126	-0.096	-0.080
County HH Income)	(0.027)	(0.026)	(0.023)	(0.021)	(0.019)	(0.027)	(0.027)	(0.024)	(0.020)	(0.019)
People all Ages	-0.003					-0.001				
in Poverty (%)	(0.001)					(0.001)				
People under 18		-0.003					-0.001			
in Poverty (%)		(0.001)					(0.001)			
Children Age 5-17			-0.003					-0.002		
in Families in Poverty (%)			(0.001)					(0.001)		
Control Variables										
Socioeconomic Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Criminality						✓	✓	✓	✓	✓
State Dummies	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year Dummies	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	1,737,499	1,737,499	1,737,499	874,529	862,970	1,737,499	1,737,499	1,737,499	874,529	862,970
Pseudo R2	0.0641	0.0641	0.0641	0.0634	0.0641	0.0641	0.0641	0.0641	0.0635	0.0654

Note: All estimates are weighted using the final weight associated to each respondent. Robust standard errors are in parentheses, clustered by county. The period covered is 2005-2010. All the columns include state and year dummies in addition to socioeconomics variables (described in the text). Columns 6, 7, 8, 9 and 10 add one index of criminality: the number of murders and nonnegligent manslaughters known to police per capita.

Table 8: Relationship Between Life Satisfaction and Income Spillovers, SMP.

Ordered Probit	Life Satisfaction (1)	Life Satisfaction (2)	Life Satisfaction (3)	Happiness (4)	Happiness (5)	Happiness (6)
Ln (HH Income)	0.212	0.169	0.208	0.180	0.117	0.176
	(0.102)	(0.102)	(0.104)	(0.130)	(0.118)	(0.130)
Ln (Median HH	-0.197		-0.207	-0.302		-0.314
Income 0-0.20 mi)	(0.144)		(0.136)	(0.147)		(0.144)
ln (Median HH		0.244	0.257		0.282	0.301
Income 0.20-0.40 mi)		(0.240)	(0.239)		(0.289)	(0.289)
Control Variables						
Socioeconomic Controls	✓	✓	✓	✓	✓	✓
Observations	323	323	323	323	323	323
Pseudo R2	0.0396	0.0395	0.0409	0.0326	0.0312	0.0343

Note: Robust standard errors are in parentheses, clustered by precinct. All the columns include socioeconomics variables (described in the text). Household income has 16 categories. The log of household income is calculated using the middle point of each category (see Appendix).

Table 9: Relationship Between Life Satisfaction and Income Spillovers, SMP.

	Ordered Probit (1)	Ordered Probit (2)	Ordered Probit (3)	Ordered Probit (4)
Life satisfaction				
Ln (HH Income)	0.221	0.237	0.203	0.238
	(0.097)	(0.106)	(0.100)	(0.106)
Ln (Median HH	-0.171	-0.148	-0.184	-0.171
Income 0-0.20 mi)	(0.166)	(0.159)	(0.174)	(0.159)
Happiness				
Ln (HH Income)	0.184	0.192	0.168	0.177
	(0.126)	(0.131)	(0.131)	(0.136)
Ln (Median HH	-0.294	-0.283	-0.315	-0.311
Income 0-0.20 mi)	(0.173)	(0.175)	(0.168)	(0.164)
Control Variables				
Socioeconomic Controls	✓	✓	✓	✓
Housing Price	✓	✓	✓	✓
Distance - Subway	✓			✓
Distance - Park		✓		✓
Distance - Library			✓	✓
Observations	323	323	323	323
Pseudo R2 (satisfaction)	0.0421	0.0446	0.0441	0.0493
Pseudo R2 (happiness)	0.0331	0.0338	0.0382	0.0386

Note: Robust standard errors are in parentheses, clustered by precinct. All the columns include socioeconomics variables (described in the text). Household income has 16 categories. The log of household income is calculated using the middle point of each category (see Appendix).

Table 10: Summary.

	County			ZIP Code		Street	
y_i	+	+	+	+	+	+	+
y^*	-	+	-	+	+	-	-
f_i		+			+		
Control Variables							
Public Goods			✓				✓
State Dummies	✓	✓	✓				
County Dummies				✓	✓		
Year Dummies	✓	✓	✓	✓	✓		
Socioeconomic Controls	✓	✓	✓	✓	✓	✓	✓
Cluster	County	County	County	ZIP Code	ZIP Code	Precinct	Precinct

Note: See Tables 2, 4, 5, 6 and 7 for a description of county-level regressions. See Tables 3, 4 and 5 for a description of ZIP code-level regressions. See Tables 8 and 9 for a description of Street-level regressions.

VIII. Appendix

A. Creation of the variable *ln(household income)*

Household income of respondents is available in categories. The number of categories is respectively 16 and 8 for the SMP and BRFSS. Our strategy for computing the "Log of Household Income" is the following. We give the middle point of the household income category to the respondents. For instance, if a respondent answered between \$50,000 and \$60,000, then \$55,000 is his household income. Top and bottom-coded categories receive a special treatment. All top-coded income are replaced by 1.5 the value of the topcode. We verify that this has no effect on our findings by using also 1, 2 and 2.5 (Appendix Tables). For the bottom-coded income category, the value was divided by 2. Once again, this does not affect the main estimates of this paper. For the BRFSS, the median household income at the county-level and the household income of respondents are in 2005 dollars. We also used in some specifications (when it is mentioned) the household income per equivalent. This measure proposed by the OECD takes into account the number of individuals in the household. We limit the number of other adults and kids to three. Household income per equivalent is a quartic in log real family income = 1 + 0.5 (other adults) + 0.3 kids.

B. List of variables coming from U.S. Census and USA Counties

When there was no information for some years, the last year available is used as a replacement. For instance, the number of violent crimes known to police in 2008 is used for the years 2009-10.

- County Median Household Income (2005-09) U.S. Census Bureau (e.g. IPE010209D)

  Household income is total money income received in a calendar year by all household members 15 years and over. Total money income is the sum of amounts reported separately for income from wages or salaries; nonfarm self-employment; farm self-employment; social security; public assistance; and all other regularly received income such as veterans' payments, pensions, unemployment compensation, and alimony. Receipts not counted as income include various "lump-sum" payments such as capital gains or inheritances. The total represents the amount of income received before deductions for personal income taxes, etc (see Source Notes and Explanations - Appendix A at http://www.census.gov/support/USACdata.html). See Appendix A.C for details on the real median household income for ZIP codes.

- Gini Coefficient, U.S. Census Bureau

- High School Graduate or Higher, 25 years old and over - percent (2005-09) U.S. Census Bureau (EDU635209D)

- Local government finances - general revenue per capita (FY 2002) U.S. Census Bureau (LOG015202D)

- Local government finances - direct general expenditures per capita (FY 2002) U.S. Census Bureau (LOG315202D)

- Local government finances - direct general expenditures for education per capita (FY 2002) U.S. Census Bureau (LOG320202D)

- Local government finances - direct general expenditures for public welfare per capita (FY 2002) U.S. Census Bureau (LOG330202D)
- Local government finances - direct general expenditures for hospitals and health per capita (FY 2002) U.S. Census Bureau (LOG340202D)
- Median value of specified owner-occupied housing units (sample in 2000) U.S. Census Bureau (HSG495200D)
- Number of Murders and Nonnegligent Manslaughters Known to Police (2005-2008) U.S. Census Bureau (e.g. CRM140208D)
- Number of Robberies Known to Police (2005-2008) U.S. Census Bureau (e.g. CRM160208D)
- Number of Violent Crimes Known to Police (2005-2008) U.S. Census Bureau (e.g. CRM110208D)
- Owner-Occupied Housing Units (in 2000) U.S. Census Bureau (HSG440200D)
- People of all ages in poverty - percent (2005-2009) U.S. Census Bureau (e.g. IPE120209D)
- People under age 18 in poverty - percent (2005-2009) U.S. Census Bureau (e.g. IPE220209D)
- Population per Square Mile (2010) U.S. Census Bureau (POP060210D)
- Related children age 5 to 17 in families in poverty - percent (2005-2009) U.S. Census Bureau (e.g. IPE320209D)
- Resident population 65 years and over, percent (2005-2009) U.S. Census Bureau (e.g. AGE775209D)
- Size of the area, Square Mile U.S. Census Bureau
- Unemployment Rate (county for the BRFSS) Bureau of Labor Statistics (as of 2011/11/08)
- Urban Population (in 2000) U.S. Census Bureau (POP110200D)

C. ZIP Code

- ZIP Code Median Household Income (1999) U.S. Census Bureau
- ZIP Code Median Household Income (2011) U.S. Census Bureau

  The maximum value for the ZIP code median household income in 2011 is $250,000. We use different multiplication factors in this paper. When not specified, we multiply $250,000 by 1.5. We verify that this has no effect on our findings by using also the following factors: 1, 2 and 2.5.

- Size of ZCTAs (2010), Square Meter, GIS: projection is North America Albers Equal Area Conic

The U.S. Census Bureau does not report statistics at the ZIP code-level (five-digit) since the land area covered is not always well identified. Instead, the 2000 Census and 2010 Census report statistics for ZIP Code Tabulation Areas (ZCTAs). ZCTAs are: "generalized area representations of U.S. Postal Services (USPS) ZIP code service areas. They represent the most frequently occurring five-digit ZIP code found in a given area" (www.census.gov/geo/ZCTA.html). In most cases, the ZCTA code is the same as the ZIP code for an area. The 2010 Census identifies 33,120 five-digit ZCTAs. We thus rely on the ZCTAs and obtain the median household income at the ZCTA code-level from the Census. This measure is not available yet for 2009 which means that we have to rely on the 1999 median household income.

D. Ranking

"Rank" measures the second moment of the income distribution and determines the relative position in the local income distribution. This variable is created using the variable household income and by computing the number of households for a specific reference group (i.e. by county and year in the BRFSS). The household's normalized rank in the income distribution is defined as the rank in the county / number of households in the county. This normalized rank is equal to zero for the poorest household and one for the richest household. Remember that household

income is available only in categories which explains that more than one household may have the same ranking. This explains why the richest households do, most of the times, have a value near one. The only case in which one household has a rank of one is that only one household is in the top-coded income category.

E. Weights

For all our equations in this paper using county-level data, the personal sampling weights is used to have representatives sample (at the state-level). As a specification check, we also re-scaled the weights from each cycle to sum up to one for each year since the sample size of the BRFSS increased dramatically between the waves 2006-2007. This methodology had no effect on our main estimates (available upon request). See Pfeffermann (1993) and Angrist and Pischke (2009) for a discussion on the role of sampling weights.

F. Somerville

We match the address of residence with different measures of public goods. See Heller (2011) for more details.

- Subway station addresses (www.mbta.com)
- Somerville parks, City of Somerville (Census 2009)
- Public library addresses (www.somervillepubliclibrary.org/contactus.html)
- median price of recent home sales (http://www.homeinsight.com/home-value/ma/somerville.
  as of December 1st, 2012

  HomeInsight gives a market Snapshop of recent home sales near the place of residence. We calculate the median using the sale price of the seven closest home sales.

IX. Appendix Tables

Table A.1: Descriptive Statistics: ZIP Codes, BRFSS.

State	No. of Cty	No. of Cty (Sample)	Avg. Pop. Per Cty (Sample)	Avg. Cty Area (sqm) (Sample)	No. of ZIP	No. of ZIP (Sample)	Avg. Pop. Per ZIP (Sample)	Avg. ZIP Area (sqm) (Sample)
Arizona	15	15	1305	7888	364	167	117	72.5
Maine	16	16	1659	1566	399	212	125	61.8
Ohio	88	87	424	459	1160	345	107	37.4
Rhode Island	5	5	4745	321	70	59	402	15.8
South Dakota	66	66	638	1402	381	156	188	136.8
Texas	254	107	231	1258	1866	269	92	48.7
Utah	29	25	1261	1770	278	128	246	69.2
Wyoming	23	23	1223	4830	169	81	347	45.2

Note: Only states for which we have the ZIP codes are included in this table. Columns one and two show respectively the number of counties per state and the number of counties for which data is available (our sample). Columns 3 and 4 present the average population per county and the average county area for the counties in our sample. Columns 5 to 8 do the same as columns 1-4 but using the ZIP codes instead.

Table A.2: Life Satisfaction and Income Spillovers at the County-Level, BRFSS.

Panel A: Bottom-Coded Income Category Divided by 2								
Ordered Probit	Top Multi. by 1	Top Multi. by 1.5	Top Multi. by 2	Top Multi. by 2.5	Top Multi. by 1	Top Multi. by 1.5	Top Multi. by 2	Top Multi. by 2.5
Ln (Real HH Income)	0.243	0.236	0.219	0.204	-0.016	0.064	0.105	0.115
	(0.008)	(0.007)	(0.006)	(0.005)	(0.013)	(0.014)	(0.011)	(0.009)
Ln (Median County HH Income)	-0.042	-0.067	-0.076	-0.081	0.244	0.156	0.092	0.063
	(0.015)	(0.015)	(0.015)	(0.015)	(0.019)	(0.022)	(0.020)	(0.019)
Relative Rank in County					0.860	0.627	0.462	0.393
					(0.028)	(0.036)	(0.033)	(0.029)
Panel B: Bottom-Coded Income Category Divided by 1.5								
Ordered Probit	Top Multi. by 1	Top Multi. by 1.5	Top Multi. by 2	Top Multi. by 2.5	Top Multi. by 1	Top Multi. by 1.5	Top Multi. by 2	Top Multi. by 2.5
Ln (Real HH Income)	0.297	0.256	0.233	0.214	-0.006	0.105	0.145	0.144
	(0.009)	(0.007)	(0.006)	(0.005)	(0.016)	(0.016)	(0.012)	(0.009)
Ln (Median County HH Income)	-0.049	-0.073	-0.082	-0.085	0.235	0.112	0.043	0.025
	(0.015)	(0.015)	(0.015)	(0.015)	(0.021)	(0.023)	(0.020)	(0.018)
Relative Rank in County					0.836	0.512	0.339	0.297
					(0.033)	(0.041)	(0.035)	(0.030)
Control Variables								
Socioeconomic Controls	✓	✓	✓	✓	✓	✓	✓	✓
State Dummies	✓	✓	✓	✓	✓	✓	✓	✓
Year Dummies	✓	✓	✓	✓	✓	✓	✓	✓

Note: All estimates are weighted using the final weight associated to each respondent. Robust standard errors are in parentheses, clustered by county. The period covered is 2005-2010. State and year dummies are included in addition to socioeconomic variables (described in the text). Household income has 8 categories. The log of the real household income is calculated using the middle point of each category (see Appendix). The log of the median household income come from the U.S. Census Bureau, USA Counties. Observations: 1,737,499.

Table A.3: Life Satisfaction and Income Spillovers at the ZIP Code-Level, BRFSS.

At Least 50 Respondents per ZIP Code

Panel A: Bottom-Coded Income Category Divided by 2								
Ordered Probit	Top Multi. by 1	Top Multi. by 1.5	Top Multi. by 2	Top Multi. by 2.5	Top Multi. by 1	Top Multi. by 1.5	Top Multi. by 2	Top Multi. by 2.5
Ln (Real HH Income)	0.278	0.272	0.252	0.234	0.162	0.231	0.235	0.220
	(0.006)	(0.005)	(0.004)	(0.004)	(0.012)	(0.010)	(0.009)	(0.008)
Ln (Median ZIP Code HH Income)	0.078	0.062	0.057	0.055	0.182	0.103	0.076	0.072
	(0.016)	(0.015)	(0.015)	(0.015)	(0.031)	(0.024)	(0.021)	(0.020)
Relative Rank in ZIP Code					0.377	0.143	0.066	0.058
					(0.033)	(0.032)	(0.030)	(0.029)
Panel B: Bottom-Coded Income Category Divided by 1.5								
Ordered Probit	Top Multi. by 1	Top Multi. by 1.5	Top Multi. by 2	Top Multi. by 2.5	Top Multi. by 1	Top Multi. by 1.5	Top Multi. by 2	Top Multi. by 2.5
Ln (Real HH Income)	0.309	0.295	0.268	0.246	0.207	0.282	0.272	0.245
	(0.006)	(0.005)	(0.004)	(0.004)	(0.013)	(0.010)	(0.009)	(0.008)
Ln (Median ZIP Code HH Income)	0.074	0.058	0.054	0.053	0.158	0.070	0.050	0.054
	(0.016)	(0.015)	(0.014)	(0.014)	(0.029)	(0.021)	(0.019)	(0.019)
Relative Rank in ZIP Code					0.302	0.040	-0.013	0.003
					(0.034)	(0.031)	(0.029)	(0.028)
Control Variables								
Socioeconomic Controls	✓	✓	✓	✓	✓	✓	✓	✓
County Dummies	✓	✓	✓	✓	✓	✓	✓	✓
Year Dummies	✓	✓	✓	✓	✓	✓	✓	✓

Note: Robust standard errors are in parentheses, clustered by ZIP Code. The period covered is 2005-2010 (except for Texas (2007-2010)). County and year dummies are included in addition to socioeconomic variables (described in the text). Household income has 8 categories. The log of the real household income is calculated using the middle point of each category (see Appendix). The log of the median household income come from the U.S. Census Bureau, USA Counties. Observations: 221,110.

Table A.4: Life Satisfaction and Income Spillovers at the ZIP Code-Level, BRFSS.

At Least 50 Respondents per ZIP Code per Year

Panel A: Bottom-Coded Income Category Divided by 2								
Ordered Probit	Top Multi. by 1	Top Multi. by 1.5	Top Multi. by 2	Top Multi. by 2.5	Top Multi. by 1	Top Multi. by 1.5	Top Multi. by 2	Top Multi. by 2.5
Ln (Real HH Income)	0.277	0.270	0.250	0.231	0.084	0.181	0.206	0.198
	(0.008)	(0.007)	(0.006)	(0.006)	(0.015)	(0.016)	(0.014)	(0.013)
Ln (Median ZIP Code HH Income)	0.103	0.081	0.074	0.072	0.310	0.190	0.133	0.121
	(0.020)	(0.019)	(0.019)	(0.019)	(0.029)	(0.030)	(0.030)	(0.029)
Relative Rank in ZIP Code					0.613	0.304	0.164	0.137
					(0.044)	(0.054)	(0.054)	(0.052)
Panel B: Bottom-Coded Income Category Divided by 1.5								
Ordered Probit	Top Multi. by 1	Top Multi. by 1.5	Top Multi. by 2	Top Multi. by 2.5	Top Multi. by 1	Top Multi. by 1.5	Top Multi. by 2	Top Multi. by 2.5
Ln (Real HH Income)	0.307	0.270	0.265	0.243	0.119	0.181	0.257	0.231
	(0.009)	(0.007)	(0.006)	(0.006)	(0.017)	(0.016)	(0.015)	(0.013)
Ln (Median ZIP Code HH Income)	0.098	0.081	0.070	0.069	0.284	0.190	0.080	0.085
	(0.020)	(0.019)	(0.019)	(0.019)	(0.029)	(0.030)	(0.030)	(0.029)
Relative Rank in ZIP Code					0.542	0.305	0.028	0.045
					(0.048)	(0.058)	(0.055)	(0.052)
Control Variables								
Socioeconomic Controls	✓	✓	✓	✓	✓	✓	✓	✓
County Dummies	✓	✓	✓	✓	✓	✓	✓	✓
Year Dummies	✓	✓	✓	✓	✓	✓	✓	✓

Note: Robust standard errors are in parentheses, clustered by ZIP Code. The period covered is 2005-2010 (except for Texas (2007-2010)). County and year dummies are included in addition to socioeconomic variables (described in the text). Household income has 8 categories. The log of the real household income is calculated using the middle point of each category (see Appendix). The log of the median household income come from the U.S. Census Bureau, USA Counties. Observations: 119,141.

Table A.5: Somerville: 2010 Census and Sample.

Census and Community Survey		**Sample from the SMP**	
Median Household Income [2006-2010]	61,731	Median Household Income [2011]	70,000
Household Income			
[0,10 000[	7.4%		
[10 000,15 000[	5.0%		
[15 000,25 000[	8.3%		
[25 000,35 000[	7.7%		
[35 000,50 000[	13.5%		
[50 000,75 000[	17.2%		
[50 000,75 000[	14.8%		
[50 000,75 000[	16.1%		
[50 000,75 000[	5.8%		
200,000 and more	4.1%		
		Household Income	
		[0,20 000[	11.1%
		[20 000,40 000[	15.8%
		[40 000,60 000[	13.0%
		[60 000,80 000[	15.8%
		[80 000,100 000[	13.0%
		[100 000,120 000[	10.2%
		[120 000,140 000[	6.2%
		[140 000,160 000[	3.7%
		[160 000,180 000[	3.1%
		[180 000,200 000[	1.8%
		[200 000,220 000[	1.8%
		[220 000,240 000[	0.6%
		[240 000,260 000[	0.9%
		[260 000,280 000[	0.6%
		[280 000,300 000[	0.9%
		300,000 and more	1.2%
Female	50.9%	Female	58.2%
Persons 65 years and over	10.4%	Persons 65 years and over	17,5%
White	73.9%	White	81.8%
Black	6.8%	African American	3.4%
Asian	8.7%	Asian	1.8%
American Indian	0.3%	American Indian	0.3%
Employed	70.9%	Employed	61.4%
Unemployed	4.8%	Unemployed	6.1%
Observations		Observations	323

Note: See Table 1 for more descriptive statistics of the SMP.

Table A.6: Well-Being and Income Spillovers at the Street-Level, SMP.

Somerville

Ordered Probit	Life Satisfaction	Life Satisfaction	Life Satisfaction	Life Satisfaction	Happiness	Happiness	Happiness	Happiness
Panel A:								
Bottom-Coded								
Income Category	Top	Top	Top	Top	Top	Top	Top	Top
Divided by 2	Multi.	Multi.	Multi.	Multi.	Multi.	Multi.	Multi.	Multi.
	by 1	by 1.5	by 2	by 2.5	by 1	by 1.5	by 2	by 2.5
Ln (HH Income)	0.219	0.212	0.205	0.199	0.182	0.180	0.177	0.174
	(0.101)	(0.102)	(0.102)	(0.103)	(0.129)	(0.130)	(0.128)	(0.127)
Ln (Median HH	-0.197	-0.197	-0.196	-0.195	-0.300	-0.302	-0.302	-0.302
Income 0-0.20)	(0.143)	(0.144)	(0.144)	(0.145)	(0.146)	(0.147)	(0.148)	(0.148)
Panel B:								
Bottom-Coded								
Income Category	Top	Top	Top	Top	Top	Top	Top	Top
Divided by 1.5	Multi.	Multi.	Multi.	Multi.	Multi.	Multi.	Multi.	Multi.
	by 1	by 1.5	by 2	by 2.5	by 1	by 1.5	by 2	by 2.5
Ln (HH Income)	0.227	0.218	0.209	0.202	0.195	0.191	0.187	0.183
	(0.118)	(0.119)	(0.118)	(0.118)	(0.146)	(0.146)	(0.144)	(0.142)
Ln (Median	-0.202	-0.201	-0.199	-0.197	-0.329	-0.330	-0.330	-0.330
Income 0-0.20)	(0.158)	(0.159)	(0.159)	(0.159)	(0.156)	(0.157)	(0.158)	(0.158)
Control Variables								
Socioeconomic Controls	✓	✓	✓	✓	✓	✓	✓	✓
Observations	323	323	323	323	323	323	323	323

Note: Robust standard errors are in parentheses, clustered by precinct. Household income has 16 categories. The log of household income is calculated using the middle point of each category (see Appendix).